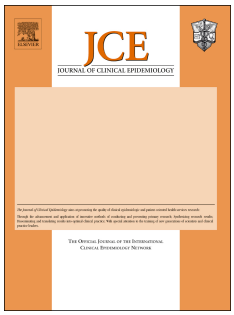
Journal Pre-proof



In network meta-analysis most of the information comes from indirect evidence: empirical study

Theodoros Papakonstantinou, Adriani Nikolakopoulou, Matthias Egger, Georgia Salanti

PII: S0895-4356(19)31016-9

DOI: https://doi.org/10.1016/j.jclinepi.2020.04.009

Reference: JCE 10122

To appear in: *Journal of Clinical Epidemiology*

Received Date: 12 November 2019

Revised Date: 23 March 2020

Accepted Date: 9 April 2020

Please cite this article as: Papakonstantinou T, Nikolakopoulou A, Egger M, Salanti G, In network meta-analysis most of the information comes from indirect evidence: empirical study, *Journal of Clinical Epidemiology* (2020), doi: https://doi.org/10.1016/j.jclinepi.2020.04.009.

In network meta-analysis most of the information comes from indirect evidence: empirical study

Theodoros Papakonstantinou[1*], Adriani Nikolakopoulou[1], Matthias Egger[1,2], Georgia Salanti[1]

[1] Institute of Social and Preventive Medicine, University of Bern, Switzerland

[2] Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, U.K.

[*]Correspondence to:

Theodoros Papakonstantinou
Institute of Social and Preventive Medicine
University of Bern
Mittelstrasse 43
CH-3012 Bern, Switzerland
theodoros.papakonstantinou@ispm.unibe.ch

Tel: +41 31 631 35 15

Abstract 192 words, main text 3637 words, 2 tables, 3 figures, 41 references.

Abstract

Objective: Network meta-analysis (NMA) may produce more precise estimates of treatment effects than pairwise meta-analysis. We examined the relative contribution of network paths of different lengths to estimates of treatment effects.

Study Design and Setting: We analyzed 213 published NMAs. We categorized network shapes according to the presence or absence of at least one closed loop (non-star or star network), and derived graph density, radius and diameter. We identified paths of different lengths and calculated their percentage contribution to each NMA effect estimate, based on their contribution matrix.

Results: Among the 213 NMAs included in analyses, 33% of the information came from paths of length 1 (direct evidence), 47% from paths of length 2 (indirect paths with one intermediate treatment) and 20% from paths of length 3. The contribution of paths of different lengths depended on the size of networks, presence of closed loops, graph radius, density and diameter. Longer paths contribute more as the number of treatments and loops, the graph radius and diameter increase.

Conclusion: The contribution of different paths depends on the size and structure of networks, with important implications for assessing the risk of bias and confidence in NMA results.

Keywords: network of interventions; network meta-analysis; paths of evidence; flow networks; study contribution; flow decomposition

Running title: contribution of indirect evidence

Word count abstract: 194

1 Introduction

Theoretical considerations and empirical data show that network meta-analysis (NMA) may produce more precise relative treatment effects than pairwise meta-analysis (1–4). The theoretical advantage of NMA arises from the integration of direct evidence - from studies directly comparing a particular treatment comparison - and indirect evidence from paths with one or more intermediate comparators (1,5,6). Methods for deriving the amount that each direct and indirect evidence route contributes for a given treatment comparison are required to estimate the amount of total indirect evidence in networks of clinical trials. The difficulty of answering this question, and the lack of empirical data, is because until recently, methods for deriving the amount that each path of direct and indirect evidence contributes were lacking. In a simple triangular network comparing treatments ABC, the AB treatment effect can be produced by synthesizing the estimate from the direct AB comparison with the indirect estimate via C. The inverse of the variance of the direct and the indirect effect, standardized over the sum of the inverse variances, is the relative contribution of the two sources of evidence. However, in more extensive networks with more complex structures,, it is not straightforward to identify the relative contributions of different sources.

The decomposition of estimates of treatment effects from NMA by the contributing sources of evidence is of interest for several reasons. Different paths of evidence might be associated with different characteristics, e.g. risk of bias, whose influence on the results could be critical (7,8). Moreover, indirect effects derived from longer routes require stronger assumptions: the transitivity assumption requires that treatment C is similar when it appears in AC and BC trials in terms of dose and dosing or administration route. Transitivity is violated if the treatment in question differs systematically between trials (9). Similarly, if the relative effectiveness of the interventions influences the choice of the comparison treatment, then the assumption of transitivity is violated. In other words, the distribution of effect modifiers of the relative treatment effects should be similar in AC and BC trials to obtain a valid AB indirect comparison (9).

In the present study we examined whether the contribution of indirect evidence might justify the additional assumptions required for longer paths in more extensive networks. Building on theoretical work (10–13), we recently developed a method to compute the contribution of different pieces of evidence to NMA relative treatment effects (14). Here we apply this approach to assess the contribution of paths of direct and indirect evidence in published NMAs of randomized control trials (RCTs).

2 Methods

2.1 Identification of networks and inclusion criteria

Petropoulou et al. compiled a database of 456 NMAs of RCTs including at least four different interventions published up to April 14, 2015; details about the search strategy, inclusion criteria and data extraction can be found elsewhere (15). For ease of use and reproducibility, we created a REDcap database containing the 456 NMAs and developed the R package `nmadb` to make it publicly accessible (16). In the present study, we included only NMAs with binary or continuous outcome data that allowed replication of the analyses. We excluded networks with evidence of inconsistency, defined by either a P value < 0.10 in the design-by-treatment interaction test (17,18) or by more than 15% of comparisons having a P value < 0.10 for SIDE (Separate Indirect from Direct Evidence) splitting (19).

2.2 Characteristics of the networks

For each network, we recorded the total number of studies, the total number of interventions, the direction of the outcome (beneficial or harmful), and the type of the outcome (objective, semi-objective or subjective). The type of outcome was defined according to the definitions given by Turner et al. In particular, objective outcomes refer to all-cause mortality. Semi-objective outcomes refer to cause-specific mortality, major morbidity event, composite mortality or morbidity, obstetric outcomes, internal structure, external structure, surgical device success or failure, withdrawals or drop-outs, resource use, and a hospital stay or process measures. Subjective outcomes refer to pain, mental health outcomes, dichotomous biological markers, quality of life or functioning, consumption, satisfaction with care,

general physical health, adverse events, infection or new disease, continuation or termination of the condition being treated, and composite endpoint (including at most one mortality or morbidity endpoint) (20). We derived several characteristics relevant to the network structure. In particular, we categorized the shape of the network according to the presence or absence of at least one closed loop (non-star or star network), and for each network, we derived the graph density, radius and diameter (21). Graph density shows how well connected a network is and is calculated as $\frac{2D}{T(T-1)}$ where D is the number of comparisons with direct data (at least one direct study) and T is the number of treatments in the network. The definition of the density can be understood as the number of comparisons with direct data over the number of all possible comparisons. As the minimum number of direct comparisons is $T-1$, density is always equal or larger than $2/T$. This implies that the minimum density is smaller for larger networks, as $T_1 > T_2 \Rightarrow \frac{2}{T_1} < \frac{2}{T_2}$. For example, a star network with five treatments has a density of 40%, whereas a star network of 10 treatments has a density of 20%. The eccentricity of a node is the maximum length of the shortest path from this node to any other. Graph radius is defined as the minimum eccentricity and graph diameter as the maximum eccentricity of a graph. We also recorded the degrees of freedom from the design-by-treatment interaction model. A review of metrics adapted from graph theory to report on NMA geometry can be found in (22).

2.3 Contribution of different direct and indirect paths in NMA

Each relative treatment effect in a network is estimated using direct and indirect evidence through various sources. Consider for example the network plot of a published NMA in Figure 1 (23). The network consists of 16 studies comparing four interventions for the treatment of achalasia of the esophagus: laparoscopic Heller myotomy (LHM), endoscopic balloon dilation (EBD), endoscopic botulinum toxin injection (EBTI) and open Heller myotomy (OHM). Clinical success expressed as odds ratios (OR) was the study outcome, as defined by clinical scores, manometric findings, or clinical interviews at 12 months. The summary OR from NMA comparing EBD and OHM is derived by

considering direct evidence from studies of EBD vs OHM and indirect information via LHM or via the EBD – EBTI – LHM – OHM paths. Paths can have different lengths; paths of length 1 represent direct evidence, paths of length 2 represent indirect evidence with one intermediate comparator, and so on.

The calculation of the contribution of a path to the estimation of a relative treatment effect is complex. We previously suggested a method (14) that utilizes the observation by König et al. (10), that each relative treatment effect can be transformed into a flow graph. We briefly describe the method using the example of Figure 1. NMA is performed as a two-stage regression model (6); in the first stage, we perform classic pairwise meta-analyses for all comparisons with direct evidence to obtain the summary estimates from the head to head trials. As the network plot of Figure 1 shows, direct evidence exists for all except the OHM vs EBTI comparison. In the second stage, we combine the five summary treatment effects from the direct comparisons to derive summary treatment effects for all combinations of treatments. The projection matrix that maps the direct effects on NMA effects involves the variances of the direct summary effects and depends on the network's structure. As heterogeneity is included in $\boldsymbol{V}$, the assumption of common or different heterogeneity values across comparisons impacts on the calculations of the projection matrix. Each row of the matrix refers to a single NMA summary treatment effect and columns refer to direct evidence. Elements of each row give information on the contribution of each treatment comparison with direct data to the particular NMA effect and can thus be seen as a generalization of pairwise meta-analysis weights.

Our method decomposes each row of the matrix to contributions of direct and indirect paths (14). Supplementary Material Table 1 shows the projection matrix of the network of interventions for the treatment of achalasia. We focus on the comparison EBD vs OHM; Figure 1 shows the row of the projection matrix corresponding to the NMA relative OR between drugs EBD and OHM. The particular row shows how much each direct comparison contributes to the estimation of the EBD vs OHM NMA effect and in particular, it shows that the information "flows" from EBD to OHM via three paths of evidence. This flow of evidence is illustrated in Figure 1, which shows the contribution of the direct and

the two different paths of indirect evidence in the estimation of the EBD versus OHM NMA relative treatment effect. Direct evidence contributes 55% to the estimation, the EBD – LHM – OHM path follows with 30% and the EBD – EBTI – LHM – OHM contributes 15%. Note that the comparison LHM vs OHM is involved in two paths (EBD – LHM – OHM and EBD – EBTI – LHM - OHM) and thus contributes more than the other comparisons that provide indirect evidence to the EBD vs OHM comparison.

Each row of Supplementary Material Table 1 can be transformed into a flow graph such as the graph of Figure 1 by using a general algorithm. We use the flow graphs to identify the paths that contribute information in each NMA treatment effect, and we derive their contribution. Described in detail in (14), the generalized algorithm can handle networks of any size.

In this empirical study, we address the contribution of different paths. However, it is also possible to split the contribution of paths to contributions of comparisons and studies. For the achalasia example, the contributions of comparisons and individual studies to all NMA treatment effects are given in Supplementary Material Tables 2, and 3, respectively and the general algorithm is described in Supplementary Material Text 1.

2.4 Analysis

We re-analyzed each network using `netmeta` (24) by including the effect sizes reported in the original publication in a random-effects model using a common heterogeneity parameter. We then identified the paths of different lengths (e.g. the paths of three different colors shown in Figure 1) and calculated their percentage contribution to each NMA effect. The paths' contributions were then aggregated by network and across networks. We plotted box plots for the contributions of paths according to their length. We also derived the mean contribution of paths of length between 1 and 5 and cumulative contributions of paths up to length 1 to 5, and we presented them separately according to the number of treatments in the network, the degrees of freedom, the radius, the graph density and the diameter. We then studied the relationship between the cumulative contribution of paths and the number of treatments in a

network. For direct paths (of length 1) the relationship is derived analytically as 2/(number of treatments). For indirect paths of length 2-5, we fitted restricted cubic splines, weighted according to the logarithm of the inverse of the standard errors. We also plotted the relationship between the cumulative distribution of paths and the derived characteristics relevant to the network structure (degrees of freedom from the design-by-treatment interaction model, graph density, radius and diameter).

For the calculations of contribution of paths, we programmed the package `nma-contribution` in the Haskell programming language, which is freely available (25). We provide an R markdown file to reproduce all analyses of this paper as Supplementary Material. There was no protocol or registration for this empirical study.

3 Results

3.1 Network characteristics

Out of the 456 published NMAs included in (15), 213 networks met the inclusion criteria. The full selection process for the included NMAs is presented in Supplementary Material Figure 1. About three quarters had a binary outcome (n=165, 77%) and one quarter a continuous outcome (n=48, 23%). Among NMAs with a binary outcome the odds ratio (OR) was the most popular effect measure (n=104) while the mean difference was most commonly used in NMAs with a continuous outcome (n=35). The number of treatments varied from 4 to 45 with a median of 6 (interquartile range [IQR] 5-9) and the median of included studies per network was 19 (IQR 11-33). A third of networks (68, 32%) did not have any closed loops. Most networks addressed either an objective (78, 37%) or a semi-objective (87, 41%) outcome. Median graph density, radius and diameter were 0.42, 1 and 2, respectively. Table 1 summarizes the main characteristics of the included NMAs.

(Table 1 here)

3.2 Contribution of paths of any length in NMA treatment effects

Figure 2 shows the path contributions for all networks ordered by their length. The median contribution of direct evidence (paths of length 1) was 33%. Around half of the information (46%) came from paths of length 2 (indirect paths with one intermediate treatment). The contribution of paths with lengths 3 and 4 (indirect paths with 2 and 3 intermediate comparators respectively) was also substantial. Among networks of any size, paths of length greater than 6 contributed only minimally (less than 2% each) to the estimation of NMA effects.

The contribution of paths of different lengths depended on the size of the network. Figure 3 shows the cumulative path contribution for paths of length 1 to 5 per number of treatments in the network. We excluded the largest network with 45 treatments to avoid predicting the contribution for network sizes of 29 to 45 treatments based on a single network. Figure 3 shows that the contribution of longer paths was more important for networks that include many treatments. For example in a network of 10 treatments, direct evidence and indirect evidence with one intermediate comparator (paths of length 1 and 2) contributed 70% of the information to the estimation of the relative treatment effects. To achieve the same level of contribution in a network of 22 treatments, paths up to length 3 need to be considered.

Table 2 summarizes the length of paths one has to consider to achieve a contribution of a certain level given the number of treatments included in the network. We considered four thresholds of cumulative contribution (70%, 80%, 90%, 95%). As an example, to achieve a contribution of 90% in a network with 20 treatments, we would have to consider paths of up to length 5. In contrast, for a network with eight treatments, paths of up to length 3 would be sufficient to achieve the same level of contribution. Figure 2, Figure 3 and Table 2 show the importance of the number of treatments involved in the network in the relative contribution of longer versus shorter paths. For instance, among networks of any size, one third of contribution comes from direct evidence and about 80% from direct evidence and indirect paths with one intermediate comparator (Figure 2). However, less than 50% of contribution comes from direct

evidence and indirect paths with one intermediate comparator for networks with more than 16 treatments (Figure 3).

(Table 2 here)

The degrees of freedom from design by treatment interaction models, graph radius, density and diameter also played a role in the relative contribution of longer versus shorter paths (Supplementary Material Figures 2-5). In networks with no closed loops (degrees of freedom equal to 0), direct evidence was more important compared to networks where loops existed. The relative importance of direct evidence decreased as loops, graph radius and diameter increased. Supplementary Material Figure 4 suggests that dense networks have larger contributions from direct evidence. The greater contributions could be because density is, by definition, larger in networks with a small number of treatments. In small networks, where there are fewer indirect paths than in large networks, we expect that the direct evidence will contribute more to the total results.

4 Discussion

In this study, we found that the contribution of direct and indirect paths of evidence to the estimation of NMA relative treatment effects depends on the size and the structure of the network. One of the main characteristics and advantages of NMA lies in the synthesis of direct and indirect evidence. The extent to which each piece of evidence contributes to the summary results has, however, been unknown. To our knowledge, this is the first study empirically assessing the relative contribution of paths of any length to the estimation of NMA relative treatment effects. Authors and reviewers interested in finding out from which sources of evidence the estimated NMA treatment effects originate can do so by using our freely available code (25,26); the two packages are the same in terms of functionality, but the Haskell package is superior in terms of speed. Alternatively, Figure 3 and Table 2 can give an approximation of the cumulative contribution of paths of lengths 1 to 6 according to the number of treatments in the network.

Empirical studies have investigated whether this theoretical advantage is important in practice. Re-analyzing 44 networks of interventions, we showed that in 20% of comparisons, the evidence for the superiority of one of the interventions was stronger with NMA than with pairwise meta-analysis (2). Lin et al. derived the distribution of the relative increase in precision in 40 NMAs. They concluded that the increase in precision of treatment estimates from NMA will depend on the availability of at least two studies contributing direct evidence and a network with well-connected treatments (3). Caldwell et al. examined the percentage increase in precision of a treatment comparison of interest in fictional networks of varying structure and amount of direct and indirect evidence (4). They found that, while including indirect evidence always increased precision, this increase was modest when direct evidence is strong and minimal when all indirect paths of one intermediate comparison are present. The empirical investigations provided insights into the conditions under which the gain in precision associated with NMA is substantial. While the relative increase by adding indirect evidence is well documented, it was unclear

until now what shorter, or longer paths contribute to the estimation of NMA effects, and whether this depends on the size and shape of the network.

One important implication of this study's results concerns the evaluation of the confidence in the relative treatment effects from NMA. Studies included in NMA may have a high risk of bias, e.g. due to inadequate blinding or concealment of allocation (27). The information on the paths that include such studies, and their contribution to each NMA treatment effect, is critical to judge the confidence that one can place in the results. Two systems have been proposed to evaluate the confidence in the results from NMA (CINeMA (Confidence In Network Meta-Analysis) and GRADE (Grading of Recommendations Assessment, Development and Evaluation)) (7,8,28,29); the two systems follow similar considerations in some, but not all, of their recommendations. One important difference lies in the fact that CINeMA makes use of the contribution matrix described in (14) and used in the present study. In contrast, the GRADE system somewhat simplistically focuses on the most influential path of length 2, along with the direct evidence (28,29). The results of this empirical study emphasize that the GRADE approach is problematic as it discards a large amount of information, in particular in networks with many treatments.

Our study has several limitations. We excluded networks that demonstrated evidence of inconsistency to make it less probable that NMA results, which inform decisions, are unreliable. However, we cannot exclude the possibility that inconsistency exists in some of the included networks because the power of tests of inconsistency is limited (30,31). Another limitation lies in the fact that there might be more than one way of selecting paths in each NMA treatment effect. We have elaborated on this issue in (14), where we showed that the consequence of this ambiguity does not substantially influence the results. We have not taken into account that the included NMAs may not be independent. Naudet et al. found that many NMAs "exhibit extensive overlap and potential redundancy" (32). It is not known how such dependencies between NMAs might have affected the relationship between the contributions of different paths and the size and structure of the network. Thus, readers should take into account that the same studies might have been included in different NMAs when interpreting our results.

This study is based on the collection of NMAs described in (15) and is to our knowledge the most extensive empirical study of NMAs conducted so far. This database was initially constructed (33) and subsequently updated (15) to be a useful resource to investigators planning simulations or empirical studies. The development of the R package `nmadb` that renders the re-use of the NMA outcome data possible constitutes a further step towards that aim (34). Apart from this study, several empirical projects have been undertaken (2,35–37) or planned (38) using this NMA database. Empirical studies using different collections of NMAs have also been conducted (39–41). The networks included in these studies differed due to differences in inclusion criteria and the recency of the literature search. The substantial increase in NMA publications in recent years renders the update of the NMA database challenging. Its transformation into a regularly updated registry of NMAs might be the way forward.

In conclusion, systematic reviewers need to consider whether strengthening the evidence with additional indirect evidence is worth the additional assumptions required. These considerations should take the broadness of the research question into account, and are informed by the findings of this empirical study. In summary, we showed that the amount by which longer paths contribute to the estimation of treatment effects from NMA is substantial, especially in networks with many treatments. The results of this empirical study reinforce the importance of using NMA in comparative effectiveness research and considering all paths of indirect evidence, including complex routes, when evaluating their results.

5 Funding

This work was supported by the Swiss National Science Foundation [special project funding (Grant No. 174281) to ME and project funding (Grant No. 179158) to TP, AN, and GS].

6 Declaration of interest

None.

7 Acknowledgements

We would like to thank the two reviewers, whose comments and suggestions greatly improved this paper.

8 Figure legends

Figure 1. Illustration of the method used to estimate the contribution of different paths, using the example of a published network meta-analysis [20] comparing four interventions for the treatment of achalasia of the oesophagus.

The colours illustrate the three paths involved in the estimation of EBD vs OHM along with their percentage contributions.

LHM: laparoscopic Heller myotomy, EBD: endoscopic balloon dilation, EBTI: endoscopic botulinum toxin injection, OHM: open Heller myotomy.

Figure 2. Box plots of contributions of paths for all NMA treatment effects in the 213 network meta-analyses included in the study according to their length.

Figure 3. Cumulative contribution of paths of length 1 to 5 to NMA treatment effects in the 213 network meta-analyses as a function of the number of treatments in the network.

Dots indicate the average cumulative contribution per number of treatments; their size is proportional to the logarithm of inverse of standard errors. Curves are fitted using restricted cubic splines, weighted according to the logarithm of the inverse of the standard errors.

9 References

1. Lumley T. Network meta-analysis for indirect treatment comparisons. Stat Med. 2002 Aug 30;21(16):2313–24.

2. Nikolakopoulou A, Mavridis D, Furukawa TA, Cipriani A, Tricco AC, Straus SE, et al. Living network meta-analysis compared with pairwise meta-analysis in comparative effectiveness research: empirical study. BMJ. 2018 Feb 28;360:k585.

3. Lin L, Xing A, Kofler MJ, Murad MH. Borrowing of strength from indirect evidence in 40 network meta-analyses. J Clin Epidemiol. 2019 Feb;106:41–9.

4. Caldwell DM, Dias S, Welton NJ. Extending Treatment Networks in Health Technology Assessment: How Far Should We Go? Value Health. 2015 Jul;18(5):673–81.

5. Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. Stat Methods Med Res. 2008 Jun;17(3):279–301.

6. Lu G, Welton NJ, Higgins JPT, White IR, Ades AE. Linear inference for mixed treatment comparison meta-analysis: a two-stage approach. Res Synth Methods. 2012 Sep;3(3):255.

7. Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the quality of evidence from a network meta-analysis. PloS One. 2014;9(7):e99682.

8. Nikolakopoulou A, Higgins JP, Papakonstantinou T, Chaimani A, Giovane CD, Egger M, et al. Assessing Confidence in the Results of Network Meta-Analysis (Cinema). bioRxiv. 2019 Apr 5;597047.

9. Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. Res Synth Methods. 2012 Jun;3(2):80–97.

10. König J, Krahn U, Binder H. Visualizing the flow of evidence in network meta-analysis and characterizing mixed treatment comparisons. Stat Med. 2013 Dec 30;32(30):5414–29.

11. Krahn U, Binder H, König J. A graphical tool for locating inconsistency in network meta-analyses. BMC Med Res Methodol. 2013 Mar 9;13:35.

12. Rücker G. Network meta-analysis, electrical networks and graph theory. Res Synth Methods. 2012 Dec;3(4):312–24.

13. Rücker G, Schwarzer G. Reduce dimension or reduce weights? Comparing two approaches to multi-arm studies in network meta-analysis. Stat Med. 2014 Nov 10;33(25):4353–69.

14. Papakonstantinou T, Nikolakopoulou A, Rücker G, Chaimani A, Schwarzer G, Egger M, et al. Estimating the contribution of studies in network meta-analysis: paths, flows and streams. F1000Research. 2018;7:610.

15. Petropoulou M, Nikolakopoulou A, Veroniki A-A, Rios P, Vafaei A, Zarin W, et al. Bibliographic study showed improving statistical methodology of network meta-analyses published between 1999 and 2015. J Clin Epidemiol. 2017 Feb;82:20–8.

16. Papakonstantinou T. nmadb: Network Meta-Analysis Database API [Internet]. 2019. Available from: https://CRAN.R-project.org/package=nmadb

17. Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. Res Synth Methods. 2012 Jun;3(2):98–110.

18. White IR, Barrett JK, Jackson D, Higgins JPT. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. Res Synth Methods. 2012 Jun;3(2):111–25.

19. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. Stat Med. 2010 Mar 30;29(7–8):932–44.

20. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. Int J Epidemiol. 2012 Jun;41(3):818–27.

21. Bondy JA, Murty USR. Graph theory. New York: Springer; 2010.

22. Tonin FS, Borba HH, Mendes AM, Wiens A, Fernandez-Llimos F, Pontarolo R. Description of network meta-analysis geometry: A metrics design study. PloS One. 2019;14(2):e0212650.

23. Schoenberg MB, Marx S, Kersten JF, Rösch T, Belle S, Kähler G, et al. Laparoscopic Heller Myotomy Versus Endoscopic Balloon Dilatation for the Treatment of Achalasia: A Network Meta-Analysis. Ann Surg. 2013 Dec;258(6):943–52.

24. Rücker G, Krahn U, König J, Efthimiou O, Schwarzer G. netmeta: Network Meta-Analysis using Frequentist Methods [Internet]. 2019. Available from: https://CRAN.R-project.org/package=netmeta

25. Papakonstantinou T. nma-contribution: Haskell package to calculate contribution of studies in network meta-analysis [Internet]. Available from: https://github.com/tpapak/nma-contribution [Internet]. Available from: https://github.com/tpapak/nma-contribution

26. Papakonstantinou T. flow_contribution: R package to calculate contribution of studies in network meta-analysis [Internet]. Avaialable from: https://github.com/esm-ispm-

unibe-ch/flow_contribution [Internet]. Available from: https://github.com/esm-ispm-unibe-ch/flow_contribution

27. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. [cited 2019 Aug 18]. Available from: http://handbook-5-1.cochrane.org/

28. Puhan MA, Schünemann HJ, Murad MH, Li T, Brignardello-Petersen R, Singh JA, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. BMJ. 2014 Sep 24;349:g5630.

29. Brignardello-Petersen R, Bonner A, Alexander PE, Siemieniuk RA, Furukawa TA, Rochwerg B, et al. Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. J Clin Epidemiol. 2018 Jan;93:36–44.

30. Song F, Clark A, Bachmann MO, Maas J. Simulation evaluation of statistical properties of methods for indirect and mixed treatment comparisons. BMC Med Res Methodol. 2012 Sep 12;12:138.

31. Veroniki AA, Mavridis D, Higgins JP, Salanti G. Characteristics of a loop of evidence that affect detection and estimation of inconsistency: a simulation study. BMC Med Res Methodol. 2014 Sep 19;14:106.

32. Naudet F, Schuit E, Ioannidis JPA. Overlapping network meta-analyses on the same topic: survey of published studies. Int J Epidemiol. 2017 Dec 1;46(6):1999–2008.

33. Nikolakopoulou A, Chaimani A, Veroniki AA, Vasiliadis HS, Schmid CH, Salanti G. Characteristics of networks of interventions: a description of a database of 186 published networks. PloS One. 2014;9(1):e86754.

34. Papakonstantinou T. nmadata: R package for accessing redcap database of network meta-analyses hosted by ispm university of bern [Internet]. Available from: https://github.com/esm-ispm-unibe-ch/nmadata [Internet]. Available from: https://github.com/esm-ispm-unibe-ch/nmadata

35. Veroniki AA, Vasiliadis HS, Higgins JPT, Salanti G. Evaluation of inconsistency in networks of interventions. Int J Epidemiol. 2013 Feb;42(1):332–45.

36. Chaimani A, Vasiliadis HS, Pandis N, Schmid CH, Welton NJ, Salanti G. Effects of study precision and risk of bias in networks of interventions: a network meta-epidemiological study. Int J Epidemiol. 2013 Aug;42(4):1120–31.

37. Zarin W, Veroniki AA, Nincic V, Vafaei A, Reynen E, Motiwala SS, et al. Characteristics and knowledge synthesis approach for 456 network meta-analyses: a scoping review. BMC Med. 2017 05;15(1):3.

38. Karahalios AE, Salanti G, Turner SL, Herbison GP, White IR, Veroniki AA, et al. An investigation of the impact of using different methods for network meta-analysis: a protocol for an empirical evaluation. Syst Rev. 2017 24;6(1):119.

39. Trinquart L, Attiche N, Bafeta A, Porcher R, Ravaud P. Uncertainty in Treatment Rankings: Reanalysis of Network Meta-analyses of Randomized Trials. Ann Intern Med. 2016 May 17;164(10):666.

40. Schuit E, Ioannidis JP. Network meta-analyses performed by contracting companies and commissioned by industry. Syst Rev [Internet]. 2016 Nov 25;5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5123429/

41. Créquit P, Martin-Montoya T, Attiche N, Trinquart L, Vivot A, Ravaud P. Living network meta-analysis was feasible when considering the pace of evidence generation. J Clin Epidemiol. 2019 Apr;108:10–6.

Table 1. Characteristics of 213 network meta-analyses included in the study.

Values are frequencies (percentages) for categorical and median (IQR) for continuous characteristics.

Characteristics of network meta-analyses		**Values**
General characteristics		
Number of studies included		19 (11-33)
Number of treatments compared		6 (5-9)
Network structure characteristics		-
Shape of network	None closed loop included (star network)	68 (32%)
	At least one closed loop included (non-star network)	145 (68%)
Degrees of freedom from the design by treatment interaction model		2 (0-6)
Graph density		0.42 (0.32-0.50)
Graph radius		1 (1-2)
Graph diameter		2 (2-3)
Outcome characteristics		-
Measurement	Binary	165 (77%)
	Continuous	48 (23%)
Effect measure	Odds Ratio	104 (49%)
	Risk Ratio	58 (27%)
	Risk Difference	3 (1.4%)
	Mean Difference	35 (16%)
	Standardized Mean Difference	13 (6.1%)
Direction of effect	Beneficial	94 (44%)
	Harmful	119 (56%)
Objectivity	Objective	78 (37%)

	Semi-objective	87 (41%)
	Subjective	48 (23%)

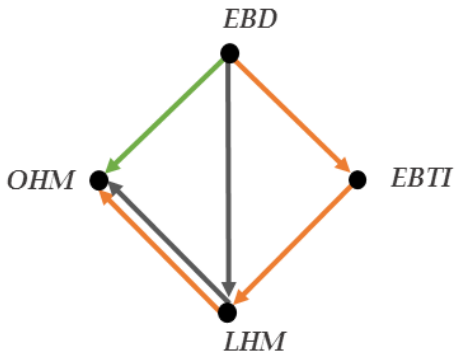
Table 2. Cumulative contribution threshold per length of path to estimates of treatment effects in the 213 network meta-analyses included in the study, according to the number of treatments in the network.

Cells refer to number of treatments included in the network that correspond to achieving the cumulative contribution indicated in the column including paths up to the length indicated in the row. NA: non applicable.

length	threshold: 0.7	threshold: 0.8	threshold: 0.9	threshold: 0.95
2	4-9	4-7	4	NA
3	10-21	8-13	5-8	4-6
4	22-29	14-29	9-19	7-13
5	NA	NA	20-29	14-23
6	NA	NA	NA	24-29

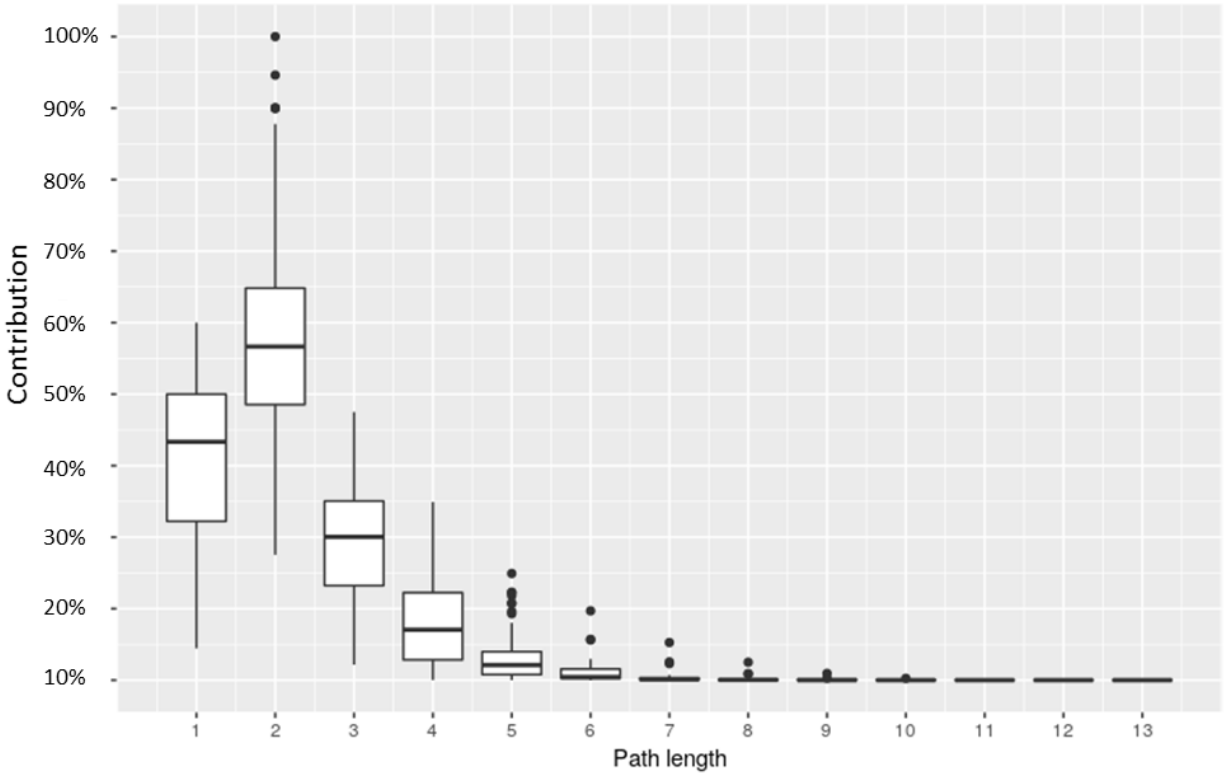
Row of projection matrix corresponding to the comparison EBD vs OHM

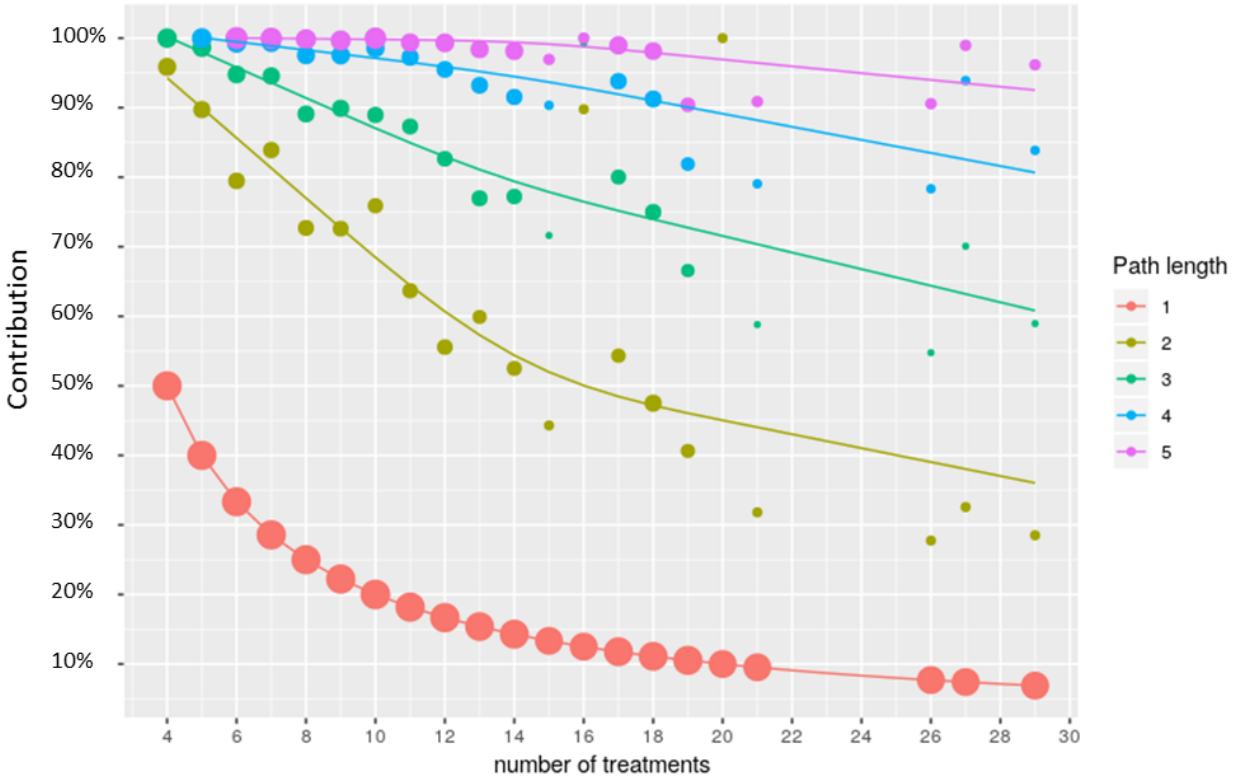
	EBD vs EBTI	EBD vs LHM	EBD vs OHM	EBTI vs LHM	LHM vs OHM
EBD vs OHM	15	30	55	15	30+15=45



Contribution of paths

	EBD – OHM	EBD – LHM – OHM	EBD – EBTI – LHM – OHM
EBD vs OHM	55%	30%	15%


100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
Contribution
1
2
3
4
5
6
7
8
9
10
11
12
13
Path length


Contribution
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
4
6
8
10
12
14
16
18
20
22
24
26
28
30
number of treatments
Path length
1
2
3
4
5

Author contributions

Theodoros Papakonstantinou: Conceptualization, Methodology, Validation, Formal analysis, Software , Writing - Original Draft, **Adriani Nikolakopoulou**: Conceptualization, Methodology, Writing - Original Draft **Matthias Egger**: Supervision, Writing - Review & Editing **Georgia Salanti**: Conceptualization, Supervision, Writing - Review & Editing, Funding acquisition, Methodology

Declaration of interest

None.

What is new?

Key findings

- The contribution of different paths of direct and indirect evidence in network meta-analysis depends on the size and structure of networks

What this adds to what is known

- Considering long paths of evidence is very important in networks with many treatments

What is the implication

- Disregarding long paths of evidence is problematic when judging the confidence that one can place in network meta-analysis results
- Authors and reviewers can use our freely available R package to examine contributions of different paths of evidence